





The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09941947

Source: 0FF

Date Processed by STIC: 09 19 200

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,

2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216. PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax) PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER</u> <u>VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

## Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address: http://www.uspto.gov/web/offices/pac/checker



Patentln 2.0

"bug"

\_Misuse of n

## Raw Sequence Listing Error Summary

## SERIAL NUMBER: SUGGESTED CORRECTION ERROR DETECTED ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SO The numberhext at the end of each line "wrapped" down to the next line. This may occur if your file Wrapped Nucleics was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will Wrapped Aminos prevent "wrapping." The rules require that a line not exceed 72 characters in length. This includes white spaces. Invalid Line Length The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; Misaligned Amino use space characters, instead. Numbering The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please Non-ASCII ensure your subsequent submission is saved in ASCII text. Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, Variable Length. each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing. A "bug" in Patentin version 2.0 has caused the <220><223> section to be missing from amino acid Patentin 2.0 Normally, Patentin would automatically generate this section from the "bug" previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences. missing. If intentional, please insert the following lines for each skipped sequence: Skipped Sequences Sequence(s) (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (OLD RULES) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION: SEQ ID NO: X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to Include the skipped sequences. missing. If Intentional, please insert the following lines for each skipped sequence Skipped Sequences (NEW RULES) <210> sequence id number <400> sequence id number Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of n's or Xaa's Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. (NEW RULES) In <220> to <223> section, please explain location of n or Xna, and which residue n or Xna represents Invalid <213> Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Response is Artificial Sequence Usc of <220> missing the <220> "Feature" and associated numeric identifiers and responses Sequence(s)

AMC/MH - Biotechnology Systems Branch - 08/21/2001

Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or

Please do not use "Copy to Disk" function of Patentln version 2.0. This causes a corrupted file,

resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence

listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

in can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represer

"Unknown." Please explain source of genetic material in <220> to <223> section.

any value not specifically a nucleotide.



OIPE

RAW SEQUENCE LISTING DATE: 09/18/2001 PATENT APPLICATION: US/09/941,947 TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\1941947.raw

```
5 <110> APPLICANT: Brzostowicz, Patricia C.
              Cheng, Qiong
     7
             DiCosimo, Deana J.
                                                                     Does Not Comply
      8
             Koffas, Mattheos
                                                                Corrected Diskette Needed
     9
             Miller, Edward S. Jr.
     10
             Odom, J. Martin
                                                                     See Poor 6 of 7A
             Picataggio, Steve
    11
             Rouviere, Pierre E.
    12
     16 <120> TITLE OF INVENTION: CAROTENOID PRODUCTION FROM A SINGLE CARBON SOURCE
     20 <130> FILE REFERENCE: CL1903 US NA
C--> 23 <140> CURRENT APPLICATION NUMBER: US/09/941,947
C--> 23 <141> CURRENT FILING DATE: 2001-08-29
     23 <150> PRIOR APPLICATION NUMBER: 60/229,907
     24 <151> PRIOR FILING DATE: 2000-09-01
     26 <150> PRIOR APPLICATION NUMBER: 60/229,858
     27 <151> PRIOR FILING DATE: 2000-09-01
     29 <160> NUMBER OF SEQ ID NOS: 60
     33 <170> SOFTWARE: Microsoft Office 97
    37 <210> SEQ ID NO: 1
     39 <211> LENGTH: 1311
     41 <212> TYPE: DNA
     43 <213> ORGANISM: Methylomonas 16a
     47 <400> SEQUENCE: 1
     48 gatgtggtca catggcccta tcacttaacg gctgatattc gattttgtca ttggtttttt
                                                                               60
     50 cttaacttta acttctacac qctcatqaac aaacctaaaa aagttgcaat actgacagca
                                                                              120
     52 ggcqqcttqq cqccttqttt qaattccqca atcqqtaqtt tqatcqaacq ttataccqaa
                                                                              180
     54 atcqatccta gcatagaaat catttgctat cgcggcggtt ataaaggcct gttgctgggc
                                                                              240
                                                                              300
     56 gattettate cagtaaegge egaagtgegt aaaaaggegg gtgttetgea aegttttgge
     58 ggttctgtga tcggcaacag ccgcgtcaaa ttgaccaatg tcaaagactg cgtgaaacgc
                                                                              360
                                                                              420
     60 ggtttggtca aagagggtga agatccgcaa aaagtcgcgg ctgatcaatt ggttaaggat
     62 ggtgtcgata ttctgcacac catcggcggc gatgatacca atacggcagc agcggatttg
                                                                              480
                                                                              540
     64 gcagcattcc tggccagaaa taattacgga ctgaccgtca ttggtttacc taaaaccgtc
                                                                              600
     66 gataacgacg tatttccgat caagcaatca ctaggtgctt ggactgccgc cgagcaaggc
                                                                              660
     68 gcgcgttatt tcatgaacgt ggtggccgaa aacaacgcca acccacgcat gctgatcgta
                                                                              720
     70 cacqaagtga tgggccgtaa ctgcggctgg ctgaccgctg caaccgcgca ggaatatcgc
                                                                              780
     72 aaattactgg accgtgccga gtggttgccg gaattgggtt tgactcgtga atcttatgaa
     74 gtgcacqcqq tattcqttcc ggaaatggcq atcgacctgg aagccgaagc caagcgcctg
                                                                              840
     76 cgcgaagtga tggacaaagt cgattgcgtc aacatcttcg tttccgaagg tgccggcgtc
                                                                              900
                                                                              960
     78 gaagctateg tegeggaaat geaggeeaaa ggeeaggaag tgeegegega tgegttegge
     80 cacatcaaac tggatgcggt caaccctggt aaatggttcg gcgagcaatt cgcgcagatg
                                                                             1020
                                                                             1080
     82 ataggcgcgg aaaaaaccct ggtacaaaaa tcgggatact tcgcccgtgc ttctgcttcc
     84 aacgttgacg acatgcgttt gatcaaatcg tgcgccgact tggcggtcga gtgcgcgttc
                                                                             1140
                                                                             1200
     86 cgccgcgagt ctggcgtgat cggtcacgac gaagacaacg gcaacgtgtt gcgtgcgatc
     88 gagtttccgc gcatcaaggg cggcaaaccg ttcaatatcg acaccgactg gttcaatagc
                                                                             1260
   * 90 atgttgagcg aaatcggcca gcctaaaggc ggtaaagtcg aagtcagcca c
     93 <210> SEQ ID NO: 2
     95 <211> LENGTH: 437
```

RAW SEQUENCE LISTING DATE: 09/18/2001 PATENT APPLICATION: US/09/941,947 TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\1941947.raw

97 <	<212	> TYI	PE: 1	PRT												
99 <	<213	> OR(	GANIS	SM: 1	Methy	ylomo	onas	16a								
				NCE:												
105	Asp	Val	Val	Thr	Trp	Pro	Tyr	His	Leu	Thr	Ala	Asp	Ile	Arg	Phe	Cys
106					5					10					15	
109	His	Trp	Phe	Phe	Leu	Asn	Phe	Asn	Phe	Tyr	Thr	Leu	Met	Asn	Lys	Pro
110		-		20					25					30	-	
	Lvs	Lvs	Val	Ala	Ile	Leu	Thr	Ala	Glv	Glv	Leu	Ala	Pro	Cvs	Leu	Asn
114	-1 -		35				-	40	4	2			45	4		-
	Ser	Ala		Glv	Ser	Leu	Ile	Glu	Ara	Tvr	Thr	Glu	Ile	Asp	Pro	Ser
118		50		2		_	55		5	-1-		60				
	Ile		Ile	Ile	Cvs	Tyr	Arg	Glv	Glv	Tvr	Lvs	Glv	Leu	Leu	Leu	Glv
122					- 4 -	70	,	1	1	-1-	75	1				80
		Ser	Tvr	Pro	Val	Thr	Ala	Glu	Val	Ara		Lvs	Ala	Glv	Val	
126	p		-1-		85					90	-1-	-1-		1	95	
	Gln	Ara	Phe	Glv		Ser	Val	Tle	Glv		Ser	Ara	Va]	Lvs		Thr
130	01	3		100	2				105			5		110		
	Asn	٧al	Lvs		Cvs	Val	Lvs	Ara		Leu	Val	Lvs	Glu		Glu	Asp
134			115	LUE	-1-		-1-	120	1			-1-	125	1		
	Pro	Gln		Val	Ala	Ala	Asp		Leu	Va1	Lvs	Asp		Va]	Asp	Ile
138	110	130	~1~			1124	135	01				140	011			
	Len		Thr	Tle	Glv	Gly		Asp	Thr	Asn	Thr		Ala	Ala	Asp	Leu
	145	*****			011	150		o.F			155					160
	_	Δla	Phe	Len	Ala	Arg	Asn	Asn	ጥህጕ	Glv		Thr	Val	Tle	Glv	
146	mra	mra	1110	шеч	165	5			- 1 -	170	Dea		, 42		175	204
	Pro	Lvs	Thr	Val		Asn	Asp	Val	Phe		Tle	Lvs	Gln	Ser		Glv
150		2,0		180	-121				185			-1-		190		1
	Ala	מיד	Thr		Ala	Glu	Gln	Glv		Ara	Tvr	Phe	Met.		Val	Val
154	•		195					200					205			
	Ala	Glu	Asn	Asn	Ala	Asn	Pro	Arq	Met	Leu	Ile	Val	His	Glu	Val	Met
158		210	•				215					220				
	Glv		Asn	Cys	Gly	Trp	Leu	Thr	Ala	Ala	Thr	Ala	Gln	Glu	Tyr	Arq
	225			•	•	230					235				-	240
		Leu	Leu	Asp	Arg	Ala	Glu	Trp	Leu	Pro	Glu	Leu	Gly	Leu	Thr	Arg
166	•			•	245			-		250			•		255	•
169	Glu	Ser	Tyr	Glu	Val	His	Ala	Val	Phe	Val	Pro	Glu	Met	Ala	Ile	Asp
170			-	260					265					270		-
173	Leu	Glu	Ala	Glu	Ala	Lys	Arg	Leu	Arg	Glu	Val	Met	Asp	Lys	Val	Asp
174			275			_	_	280	_				285	_		_
177	Cys	Val	Asn	Ile	Phe	Val	Ser	Glu	Gly	Ala	Gly	Val	Glu	Ala	Ile	Val
178	-	290					295		_		_	300				
	Ala	Glu	Met	Gln	Ala	Lys	Gly	Gln	Glu	Val	Pro	Arg	Asp	Ala	Phe	Gly
	305					310	-				315	•	-			320
185	His	Ile	Lys	Leu	Asp	Ala	Val	Asn	Pro	Gly	Lys	Trp	Phe	Gly	Glu	Gln
186			-		325					330	-	-		-	335	
	Phe	Ala	Gln	Met		Gly	Ala	Glu	Lys		Leu	Val	Gln	Lys		Gly
190				340		-			345					350		-
	Tyr	Phe	Ala		Ala	Ser	Ala	Ser	Asn	Val	Asp	Asp	Met	Arg	Leu	Ile
194	•	-	355	,				360			-	•	365	_		

RAW SEQUENCE LISTING DATE: 09/18/2001 PATENT APPLICATION: US/09/941,947 TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\I941947.raw

197 198	Lys	Ser 370	Cys	Ala	Asp	Leu	Ala 375	Val	Glu	Cys	Ala	Phe 380	Arg	Arg	Glu	Ser	
	Gly 385	Val	Ile	Gly	His	Asp 390	Glu	Asp	Asn	Gly	Asn 395		Leu	Arg	Ala	Ile 400	
		Phe	Pro	Arg	Ile 405		Gly	Gly	Lys	Pro 410	Phe	Asn	Ile	Asp	Thr 415	Asp	
	Trp	Phe	Asn	Ser 420	Met	Leu	Ser	Glu	Ile 425		Gln	Pro	Lys	Gly 430	Gly	Lys	
	Val	Glu	Val 435		His												
	Z210	)> </td <td>EQ II</td> <td>סמ כ</td> <td>. 3</td> <td></td>	EQ II	סמ כ	. 3												
			ENGTI														
			PE:														
			RGANI		Metl	vlor	nonas	16	3								
			QUE			-, - 0.			•								
						ac ca	atcaa	aggaa	a ata	catga	acca	ccto	equa	cat 1	tatqo	cggto	60
	-															gtggd	
		-														gtato	
	_															cctte	
	_	_	-	_			_		-							atcta	
																gcgag	
		-	-	_	-			_								cege	
	-															gtccg	
					-						_	-	-			cctgo	
			-				_	-	_							_	
246	gtco	ggcgg	get d	cctg	gatgo	gc gc	ccgg	ccgat	cto	gtag	gatg	ccga	agad	ctg g	ggcgg	gaaato	600
											gatg	ccga	agad	ctg q	ggcgg	gaaato	600 636
248	acgo	ggc	get o ggg o EQ II	gago	gagg						gatg	ccga	agad	ctg q	ggcgg	gaaato	
248 251	acgo <210	ggc )> SI	igg (	gago NO	cgago : 4						gatg	ccga	agad	ctg q	ggcgg	gaaato	
<ul><li>248</li><li>251</li><li>253</li></ul>	acgo <210 <211	egge )> SI l> LI	ggg ( EQ II	cgago NO H: 21	cgago : 4						gatg	ccga	agad	ctg q	ggegg	gaaato	
248 251 253 255	acga <210 <211 <212	egge )> SI l> LI 2> T	ggg ( EQ II ENGTH	cgago NO H: 21 PRT	cgagg : 4 12	je eg	gegge	catt	y aaa		gatg	ecga	agad	ctg q	ggegg	gaaato	
248 251 253 255 257	acgo <210 <211 <212 <213	cggcg )> SI l> LI 2> T: 3> OI	Jgg ( EQ II ENGTH (PE:	cgago No: H: 21 PRT ISM:	cgago : 4 L2 Meth	je eg	gegge	catt	y aaa		gatg	ccga	agad	etg g	ggegg	gaaato	
248 251 253 255 257 261	acgo <210 <211 <211 <213 <400	cggcg )> SI l> LI 2> T: 3> OI 0> SI	Jgg ( EQ II ENGTH (PE: RGAN)	cgago No: H: 21 PRT ISM: NCE:	cgagg : 4 L2 Meth	aylor	nonas	catto	y aaa	aaaa					ggcgg		
248 251 253 255 257 261	acgo <210 <213 <213 <400 Glu	cggcg )> SI l> LI 2> T: 3> OI 0> SI	Jgg ( EQ II ENGTH (PE: RGAN)	cgago No: H: 21 PRT ISM: NCE:	cgagg : 4 L2 Meth	aylor	nonas	catto	y aaa	aaaa							
248 251 253 255 257 261 263 264 267	acgo <210 <213 <213 <400 Glu 1	cggcc )> SI L> LI 2> TY 3> OI 3> SI Asn	J99 C EQ II ENGTH (PE: RGANI EQUEN	egage D NO H: 21 PRT ISM: NCE: Met	gagg : 4 L2 Meth 4 Ser	ylor Val	nonas Thr	s 16a	y aaa Lys His	Glu 10	Val	Met	Thr	Thr Val	Ser	Pro	
248 251 253 255 257 261 263 264 267 268	acgo <210 <211 <211 <400 Glu 1 Val	cggcc )> SI L> LI 2> T: 3> OI 3> SI Asn Met	Jgg ( EQ II ENGTH (PE: RGANI EQUEN Thr	egage O NO: H: 21 PRT ISM: NCE: Met Val 20	egage : 4 l2 Meth 4 Ser 5 Met	ylor Val	nonas Thr	s 16a Ile Asn	Lys His 25	Glu 10 Leu	Val Glu	Met His	Thr Ala	Thr Val	Ser 15 Pro	Pro Leu	
248 251 253 255 257 261 263 264 267 268	acgo <210 <211 <211 <400 Glu 1 Val	cggcc )> SI L> LI 2> T: 3> OI 3> SI Asn Met	Jgg ( EQ II ENGTH (PE: RGANI EQUEN Thr	egage O NO: H: 21 PRT ISM: NCE: Met Val 20	egage : 4 l2 Meth 4 Ser 5 Met	ylor Val	nonas Thr	s 16a Ile Asn	Lys His 25	Glu 10 Leu	Val Glu	Met His	Thr Ala	Thr Val	Ser 15	Pro Leu	
248 251 253 255 261 263 264 267 268 271 272	acgo <210 <211 <211 <400 Glu 1 Val	cggcg )> SI L> LI 2> TY 3> OI )> SI Asn Met	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35	PRT ISM: NCE: Met Val 20 Leu	Meth 4 Ser 5 Met	yc conylon Val Val Asp	nonas Thr Ile	s 16a Ile Asn Gly 40	Lys His 25 Leu	Glu 10 Leu Lys	Val Glu Val	Met His Leu	Thr Ala Glu 45	Thr Val 30 Ile	Ser 15 Pro	Pro Leu Leu	
248 251 253 255 257 261 263 264 267 268 271 272 275 276	acgd <210 <211 <211 <211 <400 Glu 1 Val Ala Arg	eggco )> SI  > LH  > Z 3> OI )> SI Asn Met Arg Thr 50	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro	egagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val	Meth 4 Ser 5 Met Val	yc conylor Val Val Asp	nonas Thr Ile Gly Glu 55	Ile Asn Gly 40 Cys	Lys His 25 Leu Ile	Glu 10 Leu Lys Arg	Val Glu Val Arg	Met His Leu Ile 60	Thr Ala Glu 45 Lys	Thr Val 30 Ile Ala	Ser 15 Pro Thr	Pro Leu Leu Val	
248 251 253 255 257 261 263 264 267 268 271 272 275 276	acgd <210 <211 <211 <211 <400 Glu 1 Val Ala Arg	eggco )> SI  > LH  > Z 3> OI )> SI Asn Met Arg Thr 50	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro	egagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val	Meth 4 Ser 5 Met Val	yc conylor Val Val Asp	nonas Thr Ile Gly Glu 55	Ile Asn Gly 40 Cys	Lys His 25 Leu Ile	Glu 10 Leu Lys Arg	Val Glu Val Arg	Met His Leu Ile 60	Thr Ala Glu 45 Lys	Thr Val 30 Ile Ala	Ser 15 Pro	Pro Leu Leu Val	
248 251 253 255 257 261 263 264 271 272 275 276 279 280	acgo <210 <211 <212 <213 <400 Glu 1 Val Ala Arg Pro 65	eggcops Silver Line Line Line Line Line Line Line Line	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro	cgagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile	Meth 4 Ser 5 Met Val	ylom Val Val Asp Leu Gly	nonas Thr Ile Gly Glu 55 Ala	s 16a Ile Asn Gly 40 Cys	Lys His 25 Leu Ile	Glu 10 Leu Lys Arg	Val Glu Val Arg Ile 75	Met His Leu Ile 60 Asn	Thr Ala Glu 45 Lys	Thr Val 30 Ile Ala His	Ser 15 Pro Thr Glu	Pro Leu Leu Val Leu 80	
248 251 253 255 257 261 263 264 271 272 275 276 279 280 283	acgo <210 <211 <212 <213 <400 Glu 1 Val Ala Arg Pro 65	eggcops Silver Line Line Line Line Line Line Line Line	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro	cgagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile	Meth 4 Ser 5 Met Val	ylom Val Val Asp Leu Gly	nonas Thr Ile Gly Glu 55 Ala	s 16a Ile Asn Gly 40 Cys	Lys His 25 Leu Ile	Glu 10 Leu Lys Arg	Val Glu Val Arg Ile 75	Met His Leu Ile 60 Asn	Thr Ala Glu 45 Lys	Thr Val 30 Ile Ala His	Ser 15 Pro Thr Glu Thr	Pro Leu Leu Val Leu 80	
248 251 253 255 261 263 264 267 272 275 276 279 280 283 284	acgc <210 <211 <212 <213 <400 Glu 1 Val Ala Arg Pro 65 Tyr	eggce )> SI 1> LH 2> TY 3> OI 3> OI Asn Met Arg Thr 50 Asp Gln	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro Ala Ala	cgagc D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile	Meth 4 Ser 5 Met Val Ala Val Asp 85	ylon Val Val Asp Leu Gly 70 Ala	Thr Ile Gly Glu 55 Ala Gly	s 16a Ile Asn Gly 40 Cys Gly Ala	Lys His 25 Leu Ile Thr	Glu 10 Leu Lys Arg Ile Phe 90	Val Glu Val Arg Ile 75 Ile	Met His Leu Ile 60 Asn Val	Thr Ala Glu 45 Lys Pro Ser	Thr Val 30 Ile Ala His	Ser 15 Pro Thr Glu Thr Gly 95	Pro Leu Leu Val Leu 80 Ile	
248 251 253 255 257 261 263 264 267 272 275 276 279 280 283 284 287	acgc <210 <211 <212 <213 <400 Glu 1 Val Ala Arg Pro 65 Tyr	eggce )> SI 1> LH 2> TY 3> OI 3> OI Asn Met Arg Thr 50 Asp Gln	HIGH CEQ III ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro Ala Ala	cgagd D NO H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile Ile	Meth 4 Ser 5 Met Val Ala Val Asp 85	ylon Val Val Asp Leu Gly 70 Ala	Thr Ile Gly Glu 55 Ala Gly	s 16a Ile Asn Gly 40 Cys Gly Ala	Lys His 25 Leu Ile Thr Glu Leu	Glu 10 Leu Lys Arg Ile Phe 90	Val Glu Val Arg Ile 75 Ile	Met His Leu Ile 60 Asn Val	Thr Ala Glu 45 Lys Pro Ser	Thr Val 30 Ile Ala His Pro	Ser 15 Pro Thr Glu Thr	Pro Leu Leu Val Leu 80 Ile	
248 251 253 255 257 261 263 264 267 272 275 276 279 280 283 284 287 288	acgo <210 <211 <211 <211 <400 Glu 1 Val Ala Arg Pro 65 Tyr Thr	eggco SI L> LI 2> TY 3> OI D> SI Asn Met Arg Thr 50 Asp Gln	Igg (EQ II ENGTH (PE: RGANI Thr Pro Ala 35 Pro Ala Ala Asn	PRT ISM: NCE: Met Val Leu Ile Leu 100	Meth 4 Ser 5 Met Val Ala Val Asp 85 Leu	ylon Val Val Asp Leu Gly 70 Ala	Thr Ile Gly Glu 55 Ala Gly Glu	Eatto	Lys His 25 Leu Ile Thr Glu Leu 105	Glu 10 Leu Lys Arg Ile Phe 90 Ala	Val Glu Val Arg Ile 75 Ile Ser	Met His Leu Ile 60 Asn Val	Thr Ala Glu 45 Lys Pro Ser Val	Thr Val 30 Ile Ala His Pro Pro 110	Ser 15 Pro Thr Glu Thr Gly 95 Ile	Pro Leu Val Leu 80 Ile Leu	
248 251 253 255 257 261 263 264 267 272 275 276 279 280 283 284 287 288 291	acgo <210 <211 <211 <211 <400 Glu 1 Val Ala Arg Pro 65 Tyr Thr	eggco SI L> LI 2> TY 3> OI D> SI Asn Met Arg Thr 50 Asp Gln	ggg CEQ II ENGTH (PE: RGANI EQUEN Thr Pro Ala 35 Pro Ala Ala Asn Val	PRT ISM: NCE: Met Val Leu Ile Leu 100	Meth 4 Ser 5 Met Val Ala Val Asp 85 Leu	ylon Val Val Asp Leu Gly 70 Ala	Thr Ile Gly Glu 55 Ala Gly Glu	Eatto	Lys His 25 Leu Ile Thr Glu Leu 105	Glu 10 Leu Lys Arg Ile Phe 90 Ala	Val Glu Val Arg Ile 75 Ile Ser	Met His Leu Ile 60 Asn Val	Thr Ala Glu 45 Lys Pro Ser Val Leu	Thr Val 30 Ile Ala His Pro Pro 110	Ser 15 Pro Thr Glu Thr Gly 95	Pro Leu Val Leu 80 Ile Leu	
248 251 253 255 261 263 264 267 268 271 272 275 276 283 284 287 288 291	acgo <210 <211 <211 <211 <400 Glu 1 Val Ala Arg Pro 65 Tyr Thr	eggco D> SI L> LI 2> TY 3> OI D> SI Asn Met Arg Thr 50 Asp Gln Glu	Igg CEQ II ENGTH VPE: RGANI EQUEN Thr Pro Ala 35 Pro Ala Ala Asn Val 115	cgagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile Ile Leu 100 Ile	Meth 4 Ser 5 Met Val Ala Val Asp 85 Leu	ylon Val Val Asp Leu Gly 70 Ala Asn	Thr Ile Gly Glu 55 Ala Gly Glu Ser	Eatto	Lys His 25 Leu Ile Thr Glu Leu 105 Val	Glu 10 Leu Lys Arg Ile Phe 90 Ala Met	Val Glu Val Arg Ile 75 Ile Ser Arg	Met His Leu Ile 60 Asn Val Gly Leu	Thr Ala Glu 45 Lys Pro Ser Val Leu 125	Thr Val 30 Ile Ala His Pro 110 Glu	Ser 15 Pro Thr Glu Thr Gly 95 Ile	Pro Leu Val Leu 80 Ile Leu Gly	
248 251 253 255 261 263 264 267 268 271 272 275 276 283 284 287 288 291	acgo <210 <211 <211 <211 <400 Glu 1 Val Ala Arg Pro 65 Tyr Thr	eggco D> SI L> LI 2> TY 3> OI D> SI Asn Met Arg Thr 50 Asp Gln Glu	Igg CEQ II ENGTH VPE: RGANI EQUEN Thr Pro Ala 35 Pro Ala Ala Asn Val 115	cgagd D NO: H: 21 PRT ISM: NCE: Met Val 20 Leu Val Ile Ile Leu 100 Ile	Meth 4 Ser 5 Met Val Ala Val Asp 85 Leu	ylon Val Val Asp Leu Gly 70 Ala Asn	Thr Ile Gly Glu 55 Ala Gly Glu Ser	Eatto	Lys His 25 Leu Ile Thr Glu Leu 105 Val	Glu 10 Leu Lys Arg Ile Phe 90 Ala Met	Val Glu Val Arg Ile 75 Ile Ser Arg	Met His Leu Ile 60 Asn Val Gly Leu	Thr Ala Glu 45 Lys Pro Ser Val Leu 125	Thr Val 30 Ile Ala His Pro 110 Glu	Ser 15 Pro Thr Glu Thr Gly 95 Ile	Pro Leu Val Leu 80 Ile Leu Gly	

RAW SEQUENCE LISTING DATE: 09/18/2001 PATENT APPLICATION: US/09/941,947 TIME: 10:40:16

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\I941947.raw

STATE OF THE PROPERTY OF THE P

-- 1/01/0 MIN 1 TATE TO 410/01/

```
299 Met Leu Lys Ser Leu Gly Gly Pro Leu Pro Gln Val Thr Phe Cys Pro
                        150
                                            155
303 Thr Gly Gly Val Asn Pro Lys Asn Ala Pro Glu Tyr Leu Ala Leu Lys
304
                    165
                                        170
307 Asn Val Ala Cys Val Gly Gly Ser Trp Met Ala Pro Ala Asp Leu Val
                180
                                    185
                                                        190
311 Asp Ala Glu Asp Trp Ala Glu Ile Thr Arg Arg Ala Ser Glu Ala Ala
312
           195
                                200
315 Ala Leu Lys Lys
316
       210
319 <210> SEQ ID NO: 5
321 <211> LENGTH: 1860
323 <212> TYPE: DNA
325 <213> ORGANISM: Methylomonas 16a
329 <400> SEQUENCE: 5
330 atgaaactga ccaccgacta tcccttgctt aaaaacatcc acacgccggc ggacatacgc
332 gcgctgtcca aggaccagct ccagcaactg gctgacgagg tgcgcggcta tctgacccac
                                                                          120
                                                                          180
334 acggtcagca tttccggcgg ccattttgcg gccggcctcg gcaccgtgga actgaccgtg
                                                                          240
336 gccttgcatt atgtgttcaa tacccccgtc gatcagttgg tctgggacgt gggccatcag
338 gcctatccgc acaagattct gaccggtcgc aaggagcgca tgccgaccat tcgcaccctg
                                                                          300
340 ggcqqqqtqt caqcctttcc ggcqcgggac gagagcgaat acgatgcctt cggcgtcggc
                                                                          360
                                                                          420
342 cattccagca cctcgatcag cgcggcactg ggcatggcca ttgcgtcgca gctgcgcggc
344 gaagacaaga agatggtage catcategge gaeggtteea teaceggegg catggeetat
                                                                          480
346 gaggcgatga atcatgccgg cgatgtgaat gccaacctgc tggtgatctt gaacgacaac
                                                                          540
348 gatatgtcga tctcgccgcc ggtcggggcg atgaacaatt atctgaccaa ggtgttgtcg
                                                                          600
350 agcaagtttt attcgtcggt gcgggaagag agcaagaaag ctctggccaa gatgccgtcg
                                                                          660
352 gtgtgggaac tggcgcgcaa gaccgaggaa cacgtgaagg gcatgatcgt gcccggtacc
                                                                          720
354 ttgttcgagg aattgggctt caattatttc ggcccgatcg acggccatga tgtcgagatg
                                                                          780
                                                                          840
356 ctggtgtcga ccctggaaaa tctgaaggat ttgaccgggc cggtattcct gcatgtggtg
358 accaagaagg gcaaaggcta tgcgccagcc gagaaagacc cgttggccta ccatggcgtg
                                                                          900
360 ccggctttcg atccgaccaa ggatttcctg cccaaggcgg cgccgtcgcc gcatccgacc
                                                                          960
                                                                         1020
362 tataccqaqq tqttcqqccq ctgqctgtqc qacatgqcgg ctcaagacga gcgcttgctg
                                                                         1080
364 ggcatcacgc cggcgatgcg cgaaggctct ggtttggtgg aattctcaca gaaatttccg
366 aatcqctatt tcqatqtcqc catcqccqaq caqcatqcgq tgaccttggc cgccggccag
                                                                         1140
                                                                         1200
368 gcctgccagg gcgccaagcc ggtggtggcg atttattcca ccttcctgca acgcggttac
370 gatcagttga tccacgacgt ggccttgcag aacttagata tgctctttgc actggatcgt
                                                                         1260
                                                                         1320
372 gccqqcttqq tcgqcccqqa tqgaccqacc catgctggcg cctttgatta cagctacatg
                                                                         1380
374 cgctgtattc cgaacatgct gatcatggct ccagccgacg agaacgagtg caggcagatg
376 ctgaccaccg gcttccaaca ccatggcccg gcttcggtgc gctatccgcg cggcaaaggg
                                                                         1440
378 cccqqqqcqq caatcqatcc gaccctgacc gcgctggaga tcggcaaggc cgaagtcaga
                                                                         1500
380 caccacggca geogratege cattetggce tggggcagea tggtcacgce tgeogtegaa
                                                                         1560
382 gccggcaagc agctgggcgc gacggtggtg aacatgcgtt tcgtcaagcc gttcgatcaa
                                                                         1620
384 gccttggtgc tggaattggc caggacgcac gatgtgttcg tcaccgtcga ggaaaacgtc
                                                                         1680
                                                                         1740
386 atcqccqqcq qcqctqqcaq tqcqatcaac accttcctqc aggcgcagaa ggtgctgatg
                                                                         1800
388 ccqqtctgca acatcqqcct qcccqaccqc ttcqtcqagc aaggtagtcg cgaggaattg
                                                                         1860
390 ctcaqcctqq tcqqcctcqa caqcaaqqqc atcctcqcca ccatcqaaca gttttqcgct
393 <210> SEQ ID NO: 6
395 <211> LENGTH: 620
397 <212> TYPE: PRT
```

0/10/01

والمنافق المناف المنافقة والمنافقة والمنافق المنافق المنافقة والمنافقة والمن

6 of 7 A

<210> SEQ ID NO 41 <211> LENGTH: 38 Errored <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: <400> SEQUENCE: 41 agcagctagc ggaggaataa accatgageg catttete 38 <210> SEQ ID NO 42 <211> LENGTH: 26 Great <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: <400> SEQUENCE: 42 gactagtcac gacctgctcg aacgac 26

Errored: When the ORGANISM field is "Artificial Sequence" an explanation is mandatory on field 223.

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.





<u> 1888 - Julius II. Januari Mariana da Araban Mariana da Araban Mariana da Araban Dalamana da Araban Mariana da Araban M</u>

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/941,947

DATE: 09/18/2001 TIME: 10:40:17

Input Set : A:\CL1903 US NA Seq Listing.txt
Output Set: N:\CRF3\09182001\1941947.raw

L:23 M:270 C: Current Application Number differs, Replaced Current Application No

L:23 M:271 C: Current Filing Date differs, Replaced Current Filing Date

G:3057 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:41

L:3061 M:258 W: Mandatory Feature missing, <220> FEATURE:

L:3061 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:

L:3071 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:42

L:3075 M:258 W: Mandatory Feature missing, <220> FEATURE:

L:3075 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:

0/10/01

US 0994194707P1



Creation date: 28-08-2003

Indexing Officer: FPLUMMER - FRANCIS PLUMMER

Team: OIPEBackFileIndexing

Dossier: 09941947

Legal Date: 19-02-2002

No.	Doccode	Number of pages	
1_	CRFL	8	

Total number of pages: 8

Remarks:

Order of re-scan issued on .....